DOCUMENT RESUME

ED 065 213 95 RC 006 271

AUTHOR	Ainsworth, C. L.; Christian, Chester C., Jr.
TITLE	Lubbock Bilingual Elementary Education Program, Title VII, Elementary and Secondary Education Act of 1965. Evaluation Report, 1970.
INSTITUTION	Lubbock Independent School District, Tex.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
REPORT NO	DPSC-97-219
PUB DATE	70
NOTE	32p.
EDRS PRICE	MF-$9.65 HC-$3.29
DESCRIPTORS	*Auditory Evaluation; *Bilingual Education; Curriculum Development; *Mexican Americans; *Preschool Education; *Program Evaluation; Testing

ABSTRACT

The objective of the Lubbock Bilingual Elementary Education Program evaluation was to determine the effectiveness of a bilingual (English and Spanish) program. Contained in this report is a general review discussing (1) the awareness of the need for a bilingual program, (2) the operational smoothness, (3) the staffing with bilingual teachers and aides, (4) the program approach, whereby all formal instruction was initially taught in Spanish, and (5) the curriculum development accomplished by the staff. Student progress was evaluated through use of the Peabody Picture Vocabulary Test, the interpretation of tape-recorded interviews, a diagnostic test for sound problems, and the Goodenough "Draw-a-man" Test. The report concludes that bilingual education does show a substantial and significant advantage over ordinary education for bilingual pupils. Careful testing of the bilingual pupils in the future should show the rate at which transfer of knowledge to the second language occurs. Another suggestion is that since normal monolingual children would be, in comparison to the bilingual students, at a disadvantage, it might be appropriate to initiate programs to make the entire educational system at least bilingual. (FF)

SCOPE OF INTEREST NOTICE
The ERIC Facility has assigned this document for processing to: RC FL
In our judgement, this document is also of interest to the clearinghouses noted to the right. Indexing should reflect their special points of view.

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

JUL 28 1970 REC'D

Evaluation Report
Lubbock Bilingual Elementary Education Program
Title VII, Elementary and Secondary Education Act of 1965

97-219
PA 95

RECEIVED
AUG 2 1972
NMSU
E.R.I.C.

ED 065213

Project Number DPSC-97-219

The first year of the Lubbock Bilingual Program was devoted to the planning, development, operation, and evaluation of a kindergarten in the Spanish and English languages. An evaluation design was developed for its various components and this review will pertain predominantly to that design.

Audit activities and an on-site visit by the U.S. Office of Education Program Manager for this program have yielded favorable reviews. Final evaluation of data for some of the objectives will of course depend upon longitudinal measures.

Before testing the performance objectives individually, and discussing subsequent modifications, some general and therefore, rather subjective review statements may be made:

Awareness

The school district personnel, university, and community are evidencing more awareness of need for improving opportunities for the Spanish speaking pupil. As an example, each of the candidates in a recent school board election responded positively to a question regarding need for more bilingual adults to interact with children.

Numerous visitors to the project as well as inquiries regarding its program, materials, and format are further evidence that awareness of bilingual education is growing. Sections related to bilingual education have been added to the local reading conference.

Operation

The bilingual kindergarten program is operating smoothly as a unit within the special projects division of the system. A cooperative relationship appears to exist between project staff and building principals and staff. The business function follows regular procedures of the district. Program advisory is provided from three primary sources: (a) parents who have attended four meetings and provided individual input, (b) a steering committee of

FILMED FROM BEST AVAILABLE COPY

006271

staff, the evaluator, the elementary education director and the reading specialist from the district, and (c) program consultants who work with the staff.

Staffing

The operation includes five classes, each staffed with a bilingual teacher and aide. A director, curriculum specialist, community liaison and secretary make up the remainder of the full time staff. University consultants in the areas of evaluation and audit provide part-time service as does a graphic artist for curriculum materials. More than a dozen consultants selected for their knowledge in specific areas have worked with pupils and staff during the year. Certain of the consulting staff had direct input into curriculum development. Particularly Dr. Faye Bumpass in language, Dr. Alex Crowder in discovery techniques and Mary Ann Vaughn in music have had impact upon method.

Approach

This kindergarten was initiated with all formal instruction in Spanish for the first semester with incidental English being gradually increased. Since the youngsters enrolled already possessed some English skills (as will be noted elsewhere), a modification to initiate brief periods of English at an earlier time, although maintaining a Spanish dominant program, is planned for the next year. This is an example of use of evaluation feedback into the system.

A kindergarten program approximating the model suggested by the Texas Education Agency which emphasizes instruction with an enriched environment was followed. Emphasis were upon the Spanish language, successful experiences for each child, and oral language development. Interest centers and small group instructional approaches were utilized.

Curriculum Development

A great deal of work has been accomplished by the staff in screening materials, in modifying materials and activities, and in developing curricular areas for the

bilingual kindergarten. A more thorough exploration of the curriculum design, around the topics of language, health and safety, number concepts, literature, art, music, science, and physical education is included in the curriculum guide now being printed.

The curriculum guide reflects input from teachers, aides, curriculum specialist, consultants, the steering committee and consultants. Separate dissemination documents for Adivinanzas (Riddles), Poesía Infantil (Nursery Rhymes), Juegos (Games), Canciones (Songs) and Finger Plays were developed from curriculum materials devised and adapted for the program.

Publication of the dissemination documents and the curriculum guide are in themselves revelatory of the screening and evaluation of both materials and methodology. The curriculum guide reflects refinement of composite field tried plans and units.

Evaluation

The evaluator, who initiated the Texas Tech University input into the program, and the steering committee had agreed upon basic evaluative approaches during the program planning. Project staff were also involved in refining approaches and developing instruments.

The philosophy of those involved in planning was in agreement with Campbell and Julian who state that "experimentation within schools must be conducted by the regular staff of the schools concerned whenever possible, expecially when findings are to be generalized to other classroom situations".[1] Project teachers were carefully instructed prior to standardized testing and before subject matter testing. Teachers and aides were involved in developing and using various instructional checklists.

---

[1]Donald Campbell and Stanley Julian, "Experimental and Quasi-Experimental Design for Research on Teaching", Handbook of Research on Teaching (N.L. Gage, Ed.) 1963, p. 191

Immediate feedback of evaluation of curriculum practices and materials was made so that guides could be developed which appeared to be most efficient for movement toward project goals.

Since data were not highly refined, statistical treatments were not applied to the measures of the first year's gain. Average and individual language gains are on file and included elsewhere in this report, however.

Review According to Evaluation Design

As guidelines were clarified for Title VII projects, an evaluation design was requested by the U.S. Office of Education near the middle of the first year of operation. For this project, it resulted in the recording, somewhat more specifically, of the activities already planned for the program.

The following course-product objectives and their assessment are provided. These data are reported according to component. Face validity was the criterion for data selection and although content and predictive validity are implied, only longitudinal studies can verify them.

Instructional Component

1.1 Objectives Notes Increased English and Spanish Vocabulary (Cognitive)

Scores on the Peabody Picture Vocabulary Test were obtained in both languages shortly after the children were enrolled. These are compared with scores on the same test repeated in May.

Comparison of individual pupils will give an indication whether the goal of a year's improvement, according to the test scales, has been obtained, and are useful for individual assistance.

Some generalizations have been drawn from the language testing and are reported below in tabular form for class

averages. While it is somewhat unrealistic to average test scores after they have been converted, it is useful in looking at gain from the program viewpoint. The mental age scores were used since they relate gain to advancing age (since a 7 month age change existed from pre to post test). It is recognized that the norming was developed for different pupils but, again, gain was the concern, not definitive placement according to standardized measures. There appears to be some justification for looking at class or section groups since some range is evident. When a wide range is evident for one class when compared with the other sub-groups, one may tend to question the test reliability, testing situation or examiner differences. In this program teachers were given the same instructions and materials and each examined his own class pre and post under similar conditions.

Average language gain, based upon the arithmetic mean for each class in the kindergarten programs from the MA readings of the Peabody Picture Vocabulary Test was:

Room		English $\bar{X}$	Gain	Spanish $\bar{X}$	Gain
A	pre	2.7		2.9	
	post	3.2	7 mos.	4.2	15 mos.
B	pre	3.5		4.0	
	post	4.4	11 mos.	7.5	41 mos.
C	pre	3.0		3.3	
	post	3.8	8 mos.	5.0	21 mos.
D	pre	3.2		3.2	
	post	3.9	7 mos.	3.9	7 mos.
E	pre	2.8		2.9	
	post	3.7	11 mos.	4.4	17 mos.

Generalizing from group averages it is noted that at entering a slightly higher level of Spanish or English was present. At the end of the kindergarten year (the program having been initiated in Spanish and with continuing dominance of Spanish in the daily teaching activity) more gain was revealed in Spanish but English gain was evident. The goal of a year's growth in

vocabulary was met or exceeded in English by 37 of the 78 pupils for whom data was available on both pre and post tests while 56 of the 76 pupils met or exceed arbitrarily set goals in Spanish.

This is essentially baseline information developed from the first year of operation and contains several limitations. No comparative data seems to be available for similar populations in a kindergarten program against which placement or gain could be compared. The lack of knowledge as to appropriate expectations in terms of gain as measured by the available instruments and the relationship of measures to teaching materials and procedures are not inconsequential. These limitations do not detract, however, from the usefulness of the data for this program as it proceeds.

1.2 Objective Notes Improved Speech Performance (Cognitive)

Samples of children's speech were taken on tape early in the year. Comparisons were attempted with sample post tapes for:

- sentence length
- vocabulary
- sentence pattern
- accuracy of pictoral description

The lack of standardization, both in technical terms of quality of the recording, similar stimuli and time for response as well as in use of different pictures for the primary stimulus vehicle, made generalization somewhat difficult. It has, however, prompted procedural changes for acquisition of similar data next year and has thus provided valuable feedback.

It was noted from the items that a general increase in fluency appeared to exist at the end of the year over the beginning. Counts of sentence length or number of words per response did not yield discernable difference, which may have been because of the structure rather than lack of development. Freedom of response in both languages was predominate. The exceptions to latter statement were for those youngsters who were not Spanish speaking at the beginning of the year. They appeared to remain

hesitant to respond and/or responded in one word or short phrase format in Spanish on the post tapes.

Teachers and staff observations and the typed responses indicate that more time was being spent in discussing events, pictures and the like and that more details were noted as the year progressed.

Other information from the tape samples relates to quality of pronunciation and construction. A rather wide range appears to exist in the speech of the children who speak Spanish at home as might be expected. It is also to be expected, yet perhaps noteworthy, that the intervention of Spanish in the school setting appeared to make little difference in the standardization of Spanish spoken. This may be because some regionality is noted in the teachers' pronunciation and word choices and/or ideas and conversation appeared to be sought over preciseness of vocabulary and construction.

Number Concept development was checked in both Spanish and English. During the first testing in Spanish in the first semester the pupils could generally recognize groups to five and reproduce them. Towards the end of the second semester the testing was administered in English with pupils recognizing and reproducing groups large enough to show continued growth.

2.1 Objective Notes Positive Self-Concept (Affective)

Early taped samples were taken of children. Trained observers have been asked to note gross similarities and differences of children, in their progress, to similar age children in other programs. A statement of an observer who had no previous contact with the program but who has had experience in observation and evaluation is included.

The evaluator has observed that the children now react generally in English or Spanish to incidental questions from visitors, not only according to usual language but in the language they sense the adult uses.

While the affective aspect may be harder to quantify it is evident that the children are participating willingly, and that attendance is good in the voluntary program. The

children appear to be happy and to interchange freely with peers and with adults in a confident manner.

When given a task within the range of their experience and background they participate readily. Samples of the art work of the children were kept at periodic intervals throughout the year. While assignment of mental age scores according to the Goodenough Draw-A-Man test has been questioned for these youngsters under these conditions by some familiar with the technique and the art development of children, a tremendous amount of maturation is evident in selected drawings. Samples have been maintained in which the child moved from the early named scribble representation to figures which are proportionate and which contain refinements, according to the scoring schedule, which indicate extensive growth.

3.1 Phoneme Control--English (Cognitive)

Baseline data were gathered in January by checklist which was to indicate problems. Comparisons were made with the same checklist in May. Inter-rater differences are noted, which indicate that the same rater must again rate the same child if any validity is to be maintained. Improved instrumentation with operative objectivity is needed for this item according to the evaluator.

The diagnostic sound checklist in English revealed errors which could be approached by structured language activity. Since the checklist appeared to be rather subjective, however, work on those sounds more commonly missed were given.

Another type of control was checked through the vocabulary systematically developed in English over an eight week period. A set of words were pretested and then tested again a week later. While recognizable rather than standard reproduction was required, improvement was noted on the post tests. Individual responses from both casual and structured groups teaching situations were taken. It was perhaps attributable to original placement of children in the groups but the more informal groups appeared to show more improvement.

Staff Development

1.1 Commitment to Bilingual Education (Affective)

A semantic differential was administered to obtain baseline data for this objective prior to the beginning of the program. Comparison made with scores at the end of the first year revealed no major changes within the faculty, although continued stimulation, inservice education, and evaluation may have contributed to a halo effect of the program. The teachers have reflected positive comments to other members of the staff and community.

2.1 Objective Notes Increase in Behaviorally Stated Goals (Cognitive)

Project teachers generally developed ideas in global terms unless urged to do otherwise by the curriculum specialist. A series of inservice activities were held regarding performance objectives. Somewhat more specificity is currently evident in curriculum although there is still room for improvement.

Materials and Curriculum Component

1. Objective Notes Curriculum Design and Field Testing (Cognitive)

Curriculum materials have been selected, developed, and used within one or more of the kindergarten classrooms in the program. Composite lesson plans and units of instructions have been developed in English and in Spanish. They are in the process of finalization and printing at this time.

Regionally appropriate Spanish language materials for kindergarten in the form of songs, games, finger plays and riddles have been collected. Booklets have been printed of selected items in Spanish for each of the five topics.

In summary, this component is progressing well, whether measured objectively or subjectively. Objectively, a

comprehensive curriculum guide will be available to assist the second year of operation. A number of modifications are indicated, to include

a. Reduction of repetition of Spanish vocabulary and concepts which are already known by the child when he enters.

b. Substitution of certain teaching materials, particularly in the language activity, will be made.

c. Omission of certain segments of some of the language materials.

d. Provision of time and setting for more individual oral sequence by children in both English and Spanish.

e. Increasing use of locally developed materials.

f. Development of more extensive and more objective checklists for classroom functions, for use by teachers and evaluators, particularly in phoneme control.

g. Introduction of English gradually in the early weeks of program.

h. Teaming of monolingual and bilingual teachers in the 1970-71 school year if necessary.

Summary

In short the review of the first year's operation yields a satisfactory report. This report would be enhanced if each reader could visit personally, or via the video tape recordings, the current kindergarten program. Enthusiastic, spontaneous responses in English and in Spanish are convincing evidence that the five classrooms of children are involved in a success oriented program which is approaching its objectives.

C. L. Ainsworth, Evaluator
Texas Tech University
June, 1970

JUL 28 REC'D

Inter-American Institute
University of Texas at El Paso
May 24, 1970

Mr. Abel Acosta, Director
Bilingual Education Program
Lubbock Public Schools
Lubbock, Texas

Dear Mr. Acosta:

I have completed the final report as Independent Educational Auditor for the Lubbock Bilingual Program, and am enclosing the original typescript.

It was necessary to work one additional day after leaving Lubbock in order to complete the report.

If an elaboration or explanation of any part of the report is desired, I shall be happy to provide it. After June 4, 1970, I may be contacted at the following address: c/o Sr. Humberto Marticorena, Tacna 707 Dpto 204, Magdalena del Mar, Lima, Peru.

I am also enclosing a copy of my airline ticket from El Paso to Lubbock and return, an expense of $54.00. Taxi fare from the hotel to the Lubbock airport was $2.95. The receipt from the hotel, for $23.50 (three days) is also enclosed. Payment may be sent to me at the Inter-American Institute.

If there is any immediate question to discuss, you may reach me at the residence of Hugh F. King, 5162 Waits Ave., Ft. Worth 76133; or J. E. Mayes, P. O. Box 413, Blossom, Texas 75416. Telephone of King residence 817, 923-1721; of the Mayes residence 214, 982-5343.

It has been a pleasure to work with you and with the Lubbock Bilingual Program, and I look forward with pleasure to seeing you again in the Fall. Please give my best regards to Dr. Ainsworth, Mr. Honey, Mrs. Rector, and others in the program.

Sincerely,

Chester C. Christian, Jr.

LUBBOCK BILINGUAL PROGRAM

FINAL REPORT OF

C. C. CHRISTIAN, JR., Ph.D., INDEPENDENT EDUCATIONAL AUDITOR

MAY 19, 1970

The conclusiveness of the present audit of the bilingual Education Program is limited by the fact that most final test results are not yet available. While these data are not absolutely essential to a judgment of the evaluation procedures in terms of the educational objectives set for the program, it would be possible to assess the appropriateness of, for example, the Peabody vocabulary ("intelligence") test and its translation if those test results could be compared in greater detail with pupil language performance in tape-recorded teacher-pupil interviews near the beginning and near the end of the school year.

While the ultimate scientific validity--in terms of generally accepted validation procedures as well as in terms of testing procedures and appropriateness for these particular groups--of the tests used is open to question, there seems to be no doubt on the part of either the auditor or the evaluator that a fairly accurate statement of the results of the program can be made through a careful interpretation of the test results, and that with the experience gained in this year's project, these same tests can be used to produce more nearly exact assessments in the future.

Sampling of the results of the Peabody Picture Vocabulary Test would indicate a much higher than normal rate of vocabulary growth (and estimated IQ on that basis) in Spanish. From a

sample of low, medium, and higher-scoring pupils, the average growth in a period of approximately seven months is shown to be approximately twenty-two months, with a range of growth from three months to four years and seven months. From this sample (seventeen pupils, or approximately eighteen percent of those tested), growth in measured IQ thus ranged from a minus 2 points (with two of the seventeen "losing" IQ in Spanish) to a plus 72 points! In the latter case, a girl (Anglo-American) was shown to develop (as tested in Spanish) from a mental age of three years and one month to a mental age of seven years and seven months; in another case in the same class a boy (Mexican-American) developed from a mental age of six years and one month to a mental age of ten years and seven months. In another class a girl (Mexican-American) developed from a mental age of three years and nine months to eight years and six months.

While the phenomenal progress of the Anglo-American girl in Spanish as measured by this test is an isolated case and can be explained at least in part by her relative lack of familiarity with the language upon initial testing, it should remind us that the common lot of the Mexican-American pupil is to suffer this initial disadvantage, and that this situation finds no remedy in conventional schooling.

While lack of standardization of the test and testing procedures may provide a partial explanation of the remarkable progress in Spanish of the Mexican-American pupils (with more than half their number progressing, according to this

measurement, at more than twice the normal expected rate), and while deprivation associated with poverty undoubtedly contributed to their low average initial scores, the possibility that this progress may be in great part real and lasting--and attributable to the appropriateness of the bilingual program--should not be overlooked.

According to this test, the rate of growth in mental age as tested in English was much more modest, with an average growth of eleven months for the seventeen pupils during the seven months of instruction--still a substantially higher rate than average or expected growth. Considering the fact that the test was standardized in English, this might be considered a more nearly valid measure, but again it may be that both measures are substantially accurate, and that the introduction of the pupils to school experience in a language with which they are familiar, combined with their earlier deprivation and consequent low starting point, led to an overall high rate of progress, with comparatively much greater progress in their native than in their second language.

There is no doubt that the conditions of testing, the testing procedures, and the tests themselves need much improvement, but neither is there doubt in the mind of the present writer that these tests are a fair, if rough, indication of the nature of the educational progress which is taking place, with some pupils responding extremely favorably to the program and with a few responding at a very low level. However, the

tests would indicate that this response differentiation is not closely correlated with the relative intelligence of the child; children who tested far below normal initially as well as those who tested above normal showed much higher than expected rates of progress.

A short interview in which pictures were described in Spanish and English alternatively was held by each teacher with each pupil in her class near the beginning and again near the end of the program year. Ideally, these interviews should have provided a check on the degree of the validity of the vocabulary tests. Although they may be of considerable value in assessing pupil progress and identifying continuing problems, standardization was so weak that correlations between the interviews and the test results are difficult to establish. In some cases there are indications of the inadequacy of the written tests. For example, a pupil of age 5 years 2 months who spoke good Spanish, fluently and readily, achieved a mental age on the Spanish version of the test of three years and ten months; although his command of English was clearly inferior to his Spanish, his mental age of the English version was four years and three months. In another case, a pupil of age six years scored with a mental age of 2 years and 7 months and three years on the English and Spanish tests respectively; his ability in both languages on both testing occasions, however, might be rated from good to very good--clearly at least two years beyond the level indicated by the Peabody test and the translation.

In contrast, a native speaker of English who had a score on the Spanish version of the test in the Fall of two years and five months (or two months less than the aforementioned pupil on the English version of the test--19 items correct as opposed to 22) did not respond at all in Spanish even after being in the program all year. Thus although the test scores showed them beginning at approximately the same level in what was for each the second language, the ability in English of the native Spanish speaker would seem from the results of the interview to be clearly years in advance of the ability of the native speaker of English in her second language.

And in contrast to the above case, a native speaker of Spanish, age 5 years and 11 months, who scored a mental age on the initial tests of 5 years and 6 months in English and 4 years and 4 months in Spanish, responded readily and fluently on the initial interview in Spanish, but did not respond at all in English.

These inconsistencies clearly point to a need, not only for a more careful standardization of procedures, but also for other types of tests which are capable of rendering more nearly exact information with respect to the language experience, capability, and rate of learning of each pupil. If budgetary and other limitations do not allow a careful assessment with respect to each pupil, it would seem advisable to devote more time to the evaluation of the abilities and progress of a

---

Summary of results of Peabody Vocabulary Test and teacher-pupil interviews in English and in Spanish for selected pupils:

Age at test #1	Estimate from Vocabulary test of mental age*		Interview results
Pupil #1			
5 years, 2 mths	3 yrs, 10 mths	Spanish	Very good
	4 yrs, 3 mths	English	Fair
Pupil #2			
6 years, 0 mths	3 yrs, 0 mths	Spanish	Good to very good
	2 yrs, 7 mths	English	Good to very good
Pupil #3 (Native speaker of English)			
5 years, 9 mths	2 yrs, 4 mths	Spanish	No response
	3 yrs, 7 mths	English	Fair
Pupil #4			
5 years, 11 mths	4 yrs, 4 mths	Spanish	Very Good
	5 yrs, 6 mths	English	No response

* It should be noted that the primary objective in using the Peabody Picture Vocabulary Test and the translation of that test into Spanish was not to determine mental age, but to indicate the extent of vocabulary development in each language during the program year. Thus it was not intended to measure the abilities of these pupils relative to those of others, but to measure the relative progress of each pupil in vocabulary development.

---

sample of pupils (either selected at random or in terms of range of problems represented) than to amass less reliable data with respect to the entire group. In the absence of standardized tests by means of which bilingual programs might be accurately evaluated, the only practical alternative would seem to be the design and implementation of a few studies in depth of the abilities and progress of selected pupils. This should of course include study of the family and neighborhood as well as the entire educational situation as it relates to each pupil. Information thus obtained would facilitate the development of instruments by means of which the progress of each pupil in extensive programs could be described with greater accuracy than is now possible.

Some of the factors influencing the interpretation of the tape recorded interviews are as follows: 1) Although interpretation was to be based in part on sentence length, most of the teachers asked questions most appropriately answered by one word, especially when an open-ended question failed to elicit a response, 2) When the pupil seemed not to understand the question asked in one language, some teachers would repeat it in the other language, 3) Somewhat extended explanations would be given by the teacher, sometimes before and sometimes after the question was answered, 4) the length of time for the entire interview was apparently limited, so that if the child did not respond almost immediately the interview was discontinued, 5) the techniques of the interview differed substantially with different teachers, 6) the interview

was often conducted in the midst of other and obviously distracting class activities, 7) recorders were used which did not have automatic volume control, those where speech was at a very low volume are almost inaudible; other technical characteristics also varied widely, with the inferior acoustic situation often distorting the recording, especially since recordings were made at the slowest possible speed.

Even with all the above and other limitations, the tape recordings are possibly the most valuable single item available for judging the level of ability of the pupils in each language and their progress during the year. An interpretation of at least a sample of the interviews by a person trained in phonology and morphology could result in valuable suggestions with respect to the linguistic aspects of the program.

It would have been valuable to have, in addition to these recordings, another series made by one person trained in interviewing and testing young children. The interviewer might have entered the classroom, spoken informally with the teacher and the pupils, invited one of the more responsive of the pupils to "look at some pictures", taken him to an office where he could be given a treat and induced to engage in a fairly well controlled conversation, producing a meaningful sample of vocabulary, sentence structures, phonetic patterns, etc. in each language. With proper handling of the situation, other more timid pupils would readily volunteer after a few interviews had taken place.

It does not seem feasible to use the presently available recorded interviews for the purposes for which they were made (study of sentence length, vocabulary, sentence pattern, accuracy of pictoral description, and standard constructions), because the proper type of responses for this type of study were not generally elicited. However, it is clear that there was in general appropriate response in the appropriate language, with only a few children mixing language elements in their response ("Es un cow," for example), and a very few not responding at all in one or the other language. It is also clear that, considering the value of these somewhat rudimentary recordings in assessing abilities and progress, a more carefully designed and executed program of recording could be of great value in establishing the degree of validity of other types of tests.

A diagnostic test for sound problems was administered by each teacher to her pupils, and a comparison of sample tests of individual pupils and recordings made by those pupils seems to show that there are some inconsistencies in teacher descriptions of pupil pronunciation. This is hardly surprising in view of the fact that linguists require intensive and extended training to become able to distinguish phonemic differences and subtle influences of one language upon another. It would seem particularly difficult for persons from a given section of the country or language group to analyze accurately the speech of others in their group (or even their own speech) in terms of the precise sounds produced. In view of these

limitations, and particularly considering the lack of professional training in the interpretation of sounds, the descriptions of pupils' pronunciation in English has been well done, with the principal deficiency being failure to observe fairly subtle influences in Spanish upon English pronunciation.

A sample of the tests would seem to indicate that there is some correlation between extent of English vocabulary and ability to pronounce English, with the combination high score probably indicating a comparatively early exposure to English. However, the tests do not seem reliable enough to show whether or not relatively greater fluency in Spanish affects accuracy in English pronunciation positively or negatively, This would be a highly desirable measure; it would be a strong argument in favor of bilingual education, for example, if extent of vocabulary and accuracy of pronunciation in Spanish were positively correlated with accuracy of pronunciation in English--and there is some indication in these tests that this is the case. A more complete and accurate evaluation would seem to require the services of a linguist specializing in English as a second language in administering the diagnostic tests. The inter-rater differences noted by the evaluator could result from any one or more of many influences, not the least important of which would be inability to distinguish certain sound differences due to cultural conditioning.

The auditor is in agreement with the evaluator that "imporved instrumentation with operative objectivity is

needed for this item."

The Goodenough "Draw-a-man" test has been administered in an attempt to make an assessment of changes in relative maturity. While the evaluator is cautious in interpreting the results, it does seem clear that some differences in beginning and end of program drawings do indicate a substantial gain in maturity by some pupils. There has apparently been no attempt to relate what seem to be the more obvious gains to changes in measured mental age and response during interviews, although it might be worthwhile to attempt such a correlation.

Measures of certain abilities with numbers, including rote counting, recongition, and reproduction, were made. This was done in Spanish near the beginning of the program, and in English near the end. There seem to be a few discrepancies in this measure, with some pupils described as able to recognize and reproduce a number greater than that to which they could count by rote. In general, however, the information seems to be reliable, and to show that most pupils were able, in English, to count to a higher number, recognize the number of a larger group of objects, and put together a larger group of objects and be aware of the total number, than they could in Spanish at the beginning of the program. This would seem to indicate a transfer of knowledge from Spanish to English and a simultaneous extension of cognitave abilities. It would perhaps have been advisable to conduct this test in both languages on both occasions, however.

It is not clear whether the limitations on student performance in these number tests were the result of a lack of emphasis on teaching number or a cut-off point in testing. In any case, accomplishment seems to have ranged from something like an ability to count to 3, recognize a group of 4, and assemble a group of three by one of the less capable pupils in this area who was near the end of the year able to count to 7, recognize 5, and reproduce 5, to the performance of one of the better pupils who at the beginning could count to 10, recognize 5, and reproduce 6 (?) at near the end of the program could count to 27, recognize 13, and reproduce 15. In the latter case--one of the minor inconsistencies--it is not clear how the pupil could be said to reproduce a larger number of objects than he could recognize.

There does seem to be a significant positive correlation between the ability to count, recognize, and reproduce numbers of objects and the extent of vocabulary, with the inference that the present writer considers correct that number abilities at this level can properly be considered an aspect of verbal performance.

The evaluation of the results of teacher aides teaching English vocabulary (in terms of the 400-word pre-school vocabulary which had had rather wide use in Texas) is not yet complete. There were two groups, one structured instruction and another more informal type, and acquisition of vocabulary

in the former group seems to have begun at a lower point and remained at a lower rate throughout the program year. The factors involved have not been analyzed, and we are not prepared therefore to comment on the results or their evaluation at this point.

A brief and simple semantic differential questionnaire was given to teachers at the beginning and the end of the program year, and seems to have been of some value in the interpretation of teacher attitudes. Attitudes toward various aspects of the program--such as working with another adult in the classroom, preparing teaching materials, etc.--were generally interpreted as positive. Informal comments have seemed to indicate that teachers what they consider extensive testing of pupils one of the least attractive aspects of the program. It would seem advisable to attempt to design more complex, subtle, and comprehensive questionnaires for the teachers so that their interpretation of the program and program results could be more completely and accurately determined--and other factors, such as their attitudes toward other Mexican-Americans, Hispanic culture, the Spanish language, etc. could be assessed.

Development and evaluation of a curriculum guide has not yet reached the point where these processes could be subject to a meaningful audit.

Materials produced by those in the program, and presently being reproduced and assembled, appear to be of excellent quality with respect to design, illustration, and--most important of all--correct Spanish usage, spelling, etc.

While there has as yet been no formal evaluation of such materials, the evaluator has participated in their development.

Although there has been no formal evaluation of the quality of Spanish used in the program, it may be appropriate to state here that the auditor has noted a consistent and concentrated effort on the part of all those who use Spanish and are involved in the program to use a version of Spanish which is generally acceptable among native speakers who do not speak English. A few words adapted from English and not acceptable or understood outside the border area have been used, however; for example, in the translation of the Peabody test, "bola" (signifying "bunch" in Spanish, never "ball") for "ball;" "cachando" (non-existent in Spanish) for "catching;" and "lector" (signifying "reader" in Spanish, never "lecturer") for "lecturer." In each case, the correct words (pelota, cogiendo, conferenciante) are included as options in the test and, in the view of this auditor, should be the only words included. In some of the interviews with pupils "vegetal" (in Spanish, an area of plants or underbrush, never "vegetable") was used rather than "legumbres," the standard word. In the view of the present writer, these vocabulary items are used as a result of a lack of knowledge of Spanish vocabulary, and it is therefore improper to "credit" them to the vocabulary of those who use them, especially in establishing an IQ or mental age. It makes no more sense in Mexico to say "Comí un vegetal" than it does in the USA to say "I walked through a vegetable."

The deep concern of those involved in the program with excellence in language usage is perhaps best demonstrated in the excellence of the materials which have been developed for use in the program. The teachers, the project director, the evaluator, and others have been involved in the development of these materials, and samples delivered to the auditor are, as has been stated, of excellent quality. There have been obvious and successful attempts to use the type of Spanish spoken by well-educated native Spanish speakers, highly literate in the language. This careful and conscientious effort has produced a set of materials which could provide a standard for all other bilingual programs involving the use of Spanish.

It is more difficult, however, to evaluate the Spanish used in the classroom by the teacher in informal situations, and the recording of interviews provides almost the only indication of this available tothe auditor. These recordings indicate a high general level of competence in the use of the language, but there are a few items which should be corrected. It might be well to use more extensive recordings of the speech of each teacher as a basis for an in-service training program where non-standard items are explained and standard items are given to the teacher for subsequent use. The point is considered by the present writer to be of central importance to every bilingual program by the present writer, and therefore to merit a somewhat extended commentary here.

If one of the objectives of the bilingual program is to improve use of Spanish on the part of the pupil, it is important to evaluate the Spanish used by the teacher. One aspect of such an evaluation which is subject to controversy is that of judging the quality of regional usage. Some linguists maintain that any regional form is "correct" even though its use may be restricted to a small community of speakers. However, it does not seem desirable to evaluate usage on this basis when it may reasonably be expected to subject the pupil to disdain and perhaps to ridicule on the part of others of his own language group at some later point in his career. To assume that this will not occur is to assume that his future contacts will be limited to his present speech community, which is to assume the failure of the educational process. Any educational process which tends to perpetuate the accent and vocabulary of a rural and/or illiterate tradition in speaking can only serve to help maintain the socio-economic position of the members of that tradition, no matter what language is involved.

With any given written language, it is the spoken version which is patterned most closely by the written version which carries the greatest prestige, and one of the important functions of education is to increase the prestige of the educated person. Improvement in language

usage serves to increase prestige, and thereby increase self-confidence and improve the self-image of the pupil. Since this is so important with respect to the Mexican-American, and is an avowed objective of the bilingual program, it seems that evaluation should stress standard language usage. Any attempt to improve the self-image by the inverse process of assigning prestige to forms associated with illiteracy is, in the long run, irrevocably doomed to failure. It is at best a short-term artificial classroom process.

The above position is taken in full recognition of the long-entrenched abuse of the "correction" process in the classrooms of the Southwest by non-native speakers of Spanish who attempt to compensate for their own inferiority in the use of the language by making personal attacks on non-textbook language use by their pupils. The present writer assumes that contact with language models associated with literacy (teachers, recordings, books, etc.) will provide the necessary correction just as it has, in the present program, provided for significant vocabulary growth.

There is no provision in this program for a systematic and objective evaluation of development of a positive self-concept. It is therefore not possible to include a reliable assessment by the auditor of the efficacy of the evaluation of this part of the program. Observations made during periods of instruction, free play, organized play, snacks, and other activities seem to indicate that most of the children participate confidently

and wholeheartedly in the program, with not more than four or five children in the entire program seeming to show a comparatively low level of response.

Although the evaluation of recordings is time-consuming and the results of such evaluation are not highly reliable, this would seem to be one of the most effective methods available for evaluating self-concept. Pupils might, for example, be asked about their families, their friends, school-mates, abilities, etc., and the results of a skillful interview should give an adequate impression of the degree of positiveness with which the pupil views himself. For most purposes, a series of three or four recordings during a school year could show whether or not there seemed to be significant improvement in the self-concept. While it might not be possible to evaluate all recordings, a random selection or a selection of interviews with pupils seeming to have special problems might illuminate some of the processes occurring within a given group.

There have been no instruments developed especially for evaluation of the effectiveness of in-service education for teachers in the program. It would seem advisable to develop at least a check-list by means of which teacher reactions to individual training sessions could be compared. It might also be advisable to ask each teacher to write a short report on what she considered most valuable in the presentation immediately after the end of each session.

There is as yet no formal evaluation of the activities of the bilingual program related to community reactions to and participation in the program, but several types of observations have been made, including parent visits, number of parents attending meetings and participating in programs, etc. In addition, the present writer has observed that the Director of the Bilingual Project seems to be able to produce detailed information on the home situation of many of the pupils; this seems highly commendable, with personal involvement being much greater than would be normally expected of the Director of such a program.

The program does employ a community liaison worker, and it would seem feasible to give this worker a few additional duties which would make it possible to evaluate in greater detail community reaction to and involvement in the bilingual program. While routine records kept by the liaison worker will be of value in the evaluation of community development, it would be possible to have a more complete and systematic evaluation than material from such records can offer by having the worker administer a few simple questionnaires to parents at different points in the program year. The evaluator has been provided with a sample questionnaire used in Austin and San Antonio to study attitudes toward various aspects of education and the school system on the part of Mexican-American pupils and parents, and this may be used to suggest questions which may be important

with respect to the Lubbock program.

One general comment in closing may be in order. One of the greatest difficulties in a process of evaluation and auditing such as has been outlined for bilingual programs is in the attempt to a series of detailed and comprehensive observations without intruding upon the work and the privacy of those carrying out the program. Some of the limitations on observations in the present program seem to be a result of going just far enough in observation and testing to obtain valid results without creating a negative reaction on the part of workers in the program. While it is inevitable, and even in some ways desirable, for the evaluator and the auditor to be considered "outsiders" by those in the program, it might be possible, to facilitate a more detailed and valid analysis of program results, to develop, especially on the part of the teachers, a greater degree of curiosity about the results of teaching in bilingual programs as compared to teaching in ordinary school situations. They might be asked, for example, to predict what will happen in the development of certain of their pupils in the following two months (or other specified period); at the end of that period, tests could indicate the validity of the teacher's prediction as well as the program's effectiveness.

In any case, the observable results of the year's operation of this program has whetted the auditor's appetite for more information, while increasing his conviction that

bilingual education does show a substantial and significant advantage over ordinary education for bilingual pupils, even when they are very weak in one of the two languages in which teaching occurs. One of the monolingual pupils seemed to respond to the bilingual aspect of the program to a minimal degree, but to show normal progress in his native language.

Although the limitations of the tests and on the testing procedures might lead us to accept the results with caution, there seems to be no doubt about the validity of the general results, which show the superiority of the native language as a medium of instruction, with the simultaneous result of at least a normal rate of growth in the second language. Careful testing of these same children in the future should show the rate at which transfer of knowledge to the second language occurs. Ideally, almost all the children in the program should continue to develop at a faster than normal rate in both languages until they attain an average mental age corresponding to their average chronological age--but in each of two languagues rather than one! Since normal monolingual children would be, in comparison, underprivileged, it might be then appropriate to initiate programs to make the entire educational system bi(at least)lingual.